DOCUMENT RESUME

ED 069 789 TM 002 275

AUTHOR	Hilton, Thomas L.; Berglund, Gosta W.
TITLE	Sex Differences in Mathematics Achievement--A Longitudinal Study.
PUB DATE	70
NOTE	20p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Achievement Tests; Changing Attitudes; *Grade 5; Growth Patterns; Interests; *Longitudinal Studies; *Mathematics; Questionnaires; *Sex Differences
IDENTIFIERS	SCAT Q; STEP Math

ABSTRACT

With the objective of investigating sex-typed interests as possible causes of differences in mathematics achievement between the sexes, the study made use of longitudinal data from the Growth Study begun at E. T. S. in 1961. Growth in mathematics achievement as measured by STEP Math and SCAT-Q was compared with changing interest patterns as reflected in certain biographical questionnaire responses. At Grade 5 there were no differences in achievement but thereafter the boys pulled ahead, while parallel differences emerged in the percentage perceiving mathematics as interesting and as likely to be helpful in earning a living. (Author)

Sex Differences in Mathematics Achievement--A Longitudinal Study

Thomas L. Hilton and Gosta W. Berglund

Educational Testing Service and University of Uppsala

With the objective of investigating sex-typed interests as possible causes of difference in mathematics achievement between the sexes, the study made use of longitudinal data from the Growth Study, begun at E. T. S. in 1961. Growth in mathematics achievement as measured by STEP Math and SCAT-Q was compared with changing interest patterns as reflected in certain biographical questionnaire responses. At Grade 5 there were no differences in achievement but thereafter the boys pulled ahead, while parallel differences emerged in the percentage perceiving mathematics as interesting and as likely to be helpful in earning a living.

SEX DIFFERENCES IN MATHEMATICS ACHIEVEMENT--

A LONGITUDINAL STUDY

Thomas L. Hilton

Educational Testing Services

and

Gösta W. Berglund

University of Uppsala

Many previous studies have examined sex differences in ability and achievement in mathematics. Tyler (1965, p. 240ff), Anastasi (1958, p. 497), and Maccoby (1966, p. 26-28) surveyed the field of sex differences in aptitude and achievement and report that girls usually do better in verbal and linguistic studies, while boys generally do better in numerical and spatial aptitudes and in tests of arithmetical reasoning. Maccoby (1966, p. 26) points out that "during grade school years, some studies show boys beginning to forge ahead on tests of 'arithmetical reasoning,' although a number of studies reveal no sex differences on this dimension at this time. Fairly consistently, however, boys excel at arithmetical reasoning in high school, and the differences are substantially in favor of men among college students and adults."

Despite the spate of descriptive studies, however, convincing findings as to possible causes of observed differences are scarce. Witkin, et al. (1962, p. 218ff), surveyed the evidence for sex differences in problem solving, and suggested that "women as a group tend toward a global field approach in their perceptual and intellectual functioning, men toward an analytical approach"; therefore, men should do better in mathematics. Milton (1957, p. 211) seeking the cause of sex differences in problem-solving skills, some of which were mathematical in nature, showed that they could be partly accounted for by differences in sex-role identification; a positive relationship existed between more masculine identification and better problem-solving achievement for both sexes. Since feminine roles are typically more verbal and less quantitative oriented and masculine roles are typically more oriented toward quantitative tasks, imitating the same sex parent will produce differential patterns of abilities in boys and girls. However Maccoby (1966, pp. 42-43) points out "sex differences in verbal ability decline during the age period when the rise of identification and differential modeling ought to increase them," and "consistent sex differences in quantitative ability do not appear until adolescence, long after the time when boys and girls have begun to prefer same-sex models." For these reasons, Maccoby (1966, p. 43) does "not believe that the identification hypothesis provides an adequate explanation of the sex differences in ability profiles."

Cary (1958, p. 260) reported that sex differences in problem solving are a function of sex differences in attitudes towards problem solving. Lindgren, et al (1964, p. 45) developed this idea to show that "attitudes favorable to situations involving the solving of problems were positively and significantly correlated to achievement" in arithmetic, while Alpert, et al (1963) showed that students' attitudes toward mathematics were linked with parents' conception of the educational goals of a school mathematics course and with the extent of mathematics education desired for the child by the parents.

A further explanation for sex differences in achievement which may be the most simple is suggested by Maccoby (1966, p. 40): "Members of each sex are encouraged in, and become interested in and proficient at, the kinds of tasks that are most relevant to the roles they fill currently or are expected to fill in the future. According to this view, boys in high school forge ahead in math because they and their parents and teachers know they may become engineers or scientists; while the girls know they are unlikely to need math in the occupations they will take up when they leave school." In other words, sex differences in mathematics achievement are direct effects of sex-typed interests. The "International Study of Achievement in Mathematics" (Husén, 1967), a study in the international project for the evaluation of educational achievement (IEA), provides evidence in support of this view.

The present study was designed to pursue further the hypothesis of sex-typed interests by means of longitudinal data. In accordance with the hypothesis the authors predicted the following:

1. Preadolescent differences in mathematics achievement would be negligible.

2. Beginning with adolescence (and the emergence of sex-typed interests) differences in achievement would appear and thereafter would widen in concert with widening differences in interests between the sexes.

Alpert, et al (1963) and Anttonen (1969) are cited by Aiken (1970, p. 590) as having "pointed to the need for longitudinal research on patterns of performance in mathematics emerging over time and on psychological variables related to these changes." Anttonen (1969, p. 471) maintains that the measurements in such studies should be taken over a period shorter than the six-year span he used in his study on mathematics attitude. Aiken, (1970, p. 591) concluded that periods of one or two years would be most satisfactory.

Method

Data. Data reported in the present study were taken from the Growth Study begun at Educational Testing Service in 1961 (Anderson, Maier, 1963; Hilton & Myers, 1967). As part of that study, the Sequential Test of Educational Progress (STEP) and the School and College Ability Test (SCAT) were given in 1961 to a nationwide sample of fifth graders and then again in 1963, 1965, and 1967, as seventh, ninth, and eleventh graders respectively. In addition, a 177-item Background and Experience Questionnaire (BEQ) was given each time, except in 1961.

The different forms of STEP, as well as those of SCAT, designed for each grade level, have been vertically equated. That is, they all have a common scale, and the results can be treated as if the same test was taken at each grade level. As for the BEQ, essentially the same form (except for a few items) was administered each time, the only changes being in the reading difficulty of the items.

Sample. The sample consisted of students who took three or four math courses both in grades 7 and 8 and in grades 9 and 10. Thus the amount of training in math was held approximately constant. This was an important control, for otherwise any observed differences between the sexes could be attributed to differential classroom exposure; at the high school level the boys enrolled in more math courses than the girls did.

A second control resulted from the use of matched longitudinal samples (Hilton & Patrick, 1970). The subjects at each level were the same, namely, the core of students who had complete data for all test administrations. This provision prevented the results from merely reflecting cohort changes, e.g., changes resulting from different school dropout rates between the boys and the girls.

Thirdly, curriculum was held roughly constant. In each analysis the samples were divided into two groups: (1) the students who in high school enrolled in the academic (college preparatory) program and (2) those who were enrolled in one of the several "nonacademic" programs, e.g., vocational-technical, home economics, business. Although the evidence is that these programs of instruction do not differ nearly as much as the academic-non-academic dichotomy would suggest (Hilton, 1971), it does separate the sample into more homogeneous groups as far as content of mathematics instruction is concerned. For the academic group the emphasis is on preparation for advanced mathematics, while for the nonacademic group the emphasis is on practical application.

There were 632 boys and 688 girls in the academic group, and 249 boys and 290 girls in the nonacademic group. Since, on a nationwide basis, the academic and nonacademic dichotomy splits the high school population into groups roughly equal in size, the numbers obtained here indicate that the complete data requirement and the requirement of taking three or four math courses resulted in the nonacademic group's being more highly selected than the academic group.

Instruments. The STEP series measures the student's ability to apply his skills to problem solving in six areas: reading, writing, listening, social studies, science and mathematics. In the present study STEP Math was used.

The SCAT series, which yields verbal (V) and quantitative (Q) scores, was designed to measure general ability to do school work. Its use in a number of studies based on Growth Study data (cf. Hilton & Myers, 1967) indicates that the "ability" it measures is a developed ability which, for all intents and purposes, can be considered achievement like the achievement measured by STEP. In the present study SCAT-Q was used.

BEQ is a special instrument developed for the Growth Study. It gathers information needed to relate academic growth to a student's experiences and activities in and out of school. Eight items of the BEQ were used in the present study.

Results

STEP Math and SCAT-Q. The arithmetic means and standard deviations are given in Tables 1 and 2, and the changes with grade are shown in Figures 1 and 2. The differences between the sexes at the various grade levels are shown in the tables and figures.

From inspection of the tables and figures the following conclusions can be drawn:

1. The arithmetic means increase with age without exception.

2. For all groups the standard deviations increase with age.[1] [Most investigations agree in showing an increase in individual differences with age (see, e.g., Anastasi, 1958)].

---

[1]This and other conclusions to follow all require the assumption that the STEP and SCAT scales have equal intervals over the whole range of each scale. Although this assumption cannot be asserted to be valid, the authors know of no empirical basis for rejecting it.

Ghiselli (1964) argues that variations in the equality of scale units are probably distributed at random throughout the range of the scale in the case of most achievement tests. If this is the case, it would be highly improbable that random differences in the equality of intervals would be mistaken for true differences. Where ceiling and floor effects exist, the inequality of intervals is not randomly distributed throughout the scale and change scores or curve slopes will most likely be distorted. However, no such effects are noted in our data.

Table 1

Means and Standard Deviations for STEP Math

by Sex and Curriculum

	Males		Females		Diff.	
	M	SD	M	SD	Male-Female	t
Academic						
Grade 5	252.2	10.8	251.4	9.6	0.8	1.39
Grade 7	267.2	11.3	265.8	10.3	1.4	2.37*
Grade 9	278.8	11.5	276.5	10.5	2.3	3.77***
Grade 11	286.7	12.0	282.0	11.8	4.7	7.12***
Nonacademic						
Grade 5	245.1	10.0	245.4	9.3	-0.3	0.36
Grade 7	257.2	11.9	256.4	11.5	0.8	0.80
Grade 9	267.7	12.3	266.2	11.2	1.5	1.47
Grade 11	275.4	14.8	270.4	15.1	5.0	3.88***

*Significant at the 5 per cent level.

**Significant at the 1 per cent level.

***Significant at the 0.1 per cent level.

Table 2

Means and Standard Deviations for SCAT-Q

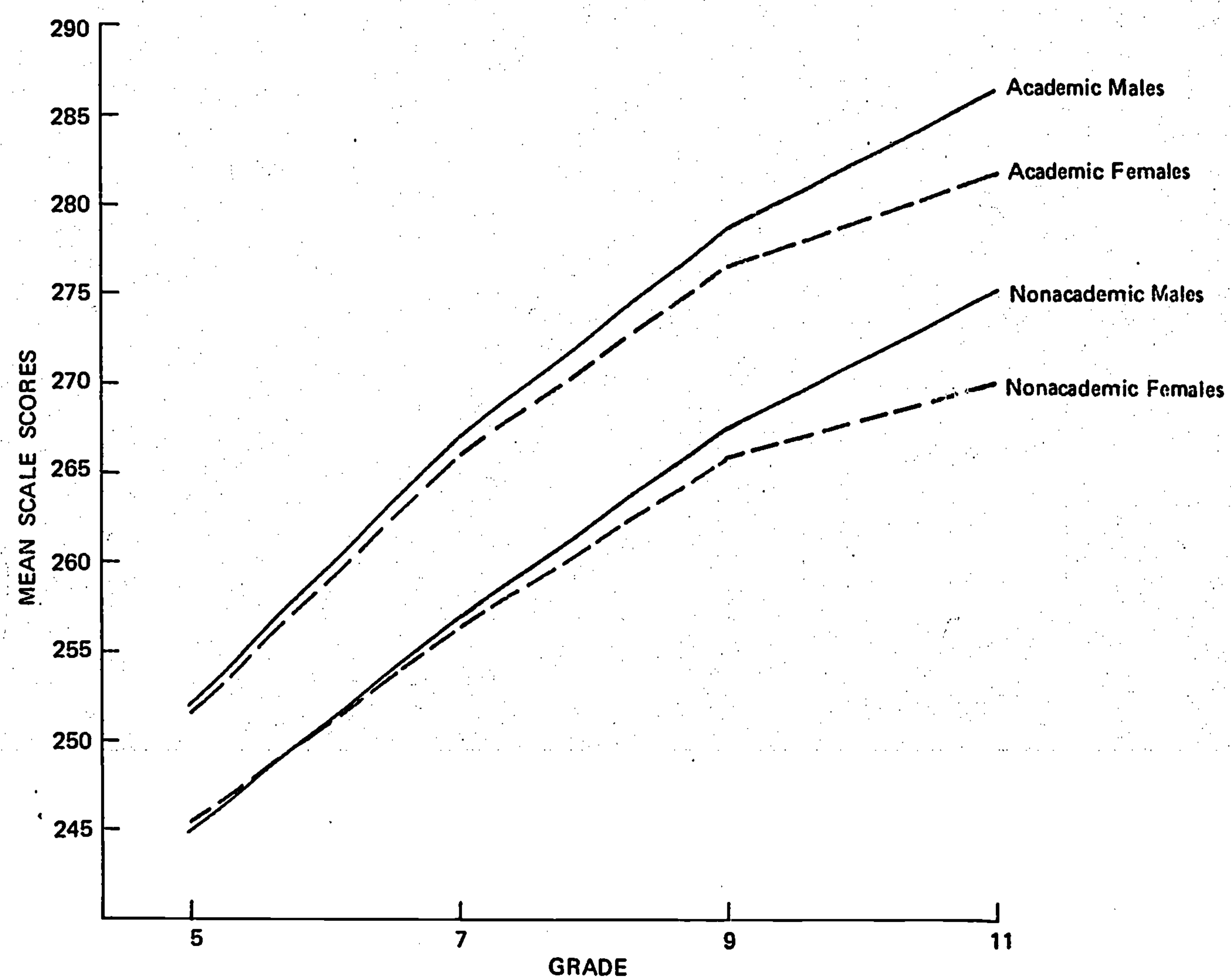
by Sex and Curriculum

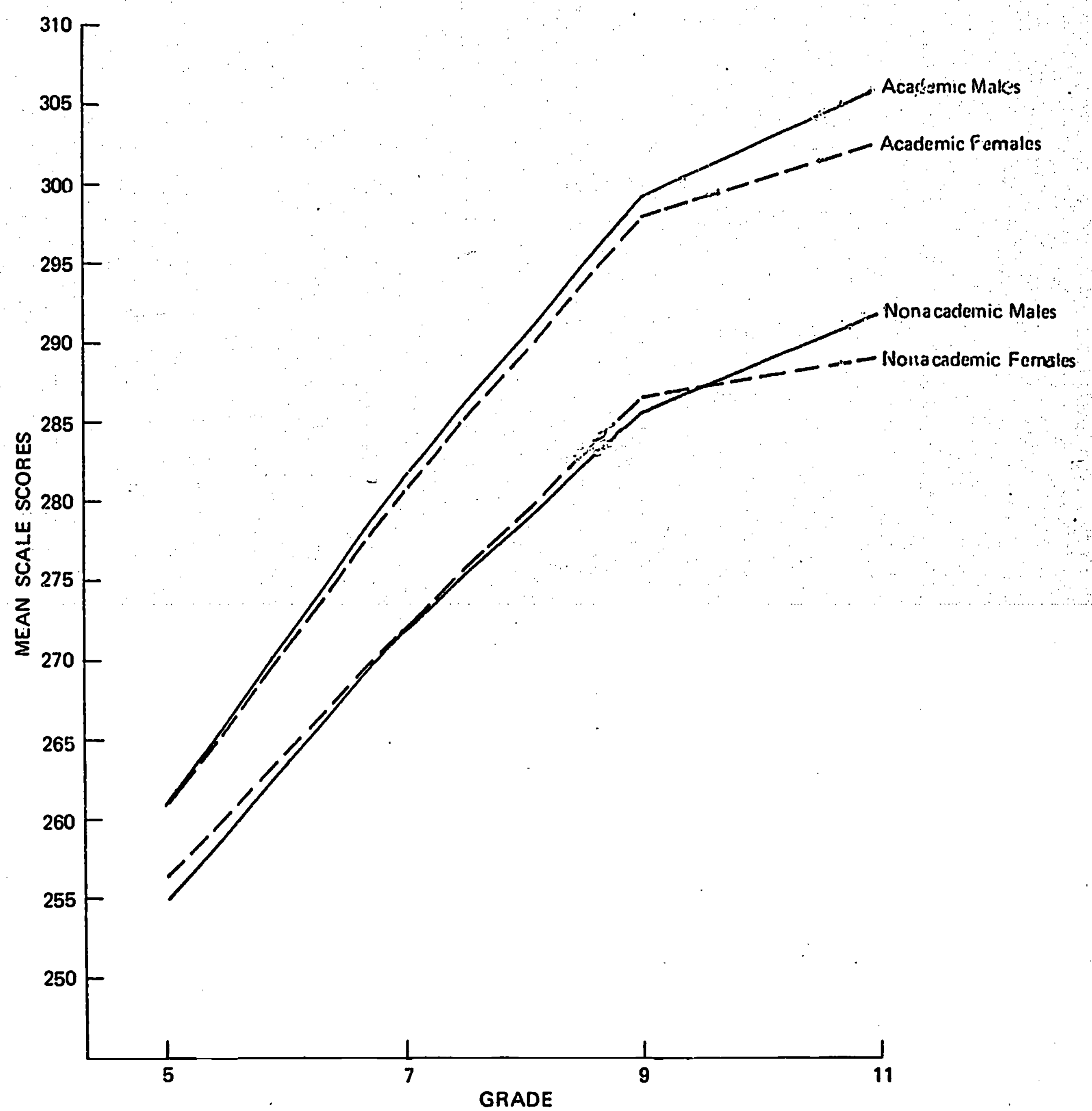
	Males		Females		Diff.	
	M	SD	M	SD	Male-Female	t
Academic						
Grade 5	260.9	8.7	260.6	8.1	0.3	0.64
Grade 7	281.7	13.8	280.8	11.7	0.9	1.27
Grade 9	299.2	14.7	298.1	13.6	1.1	1.39
Grade 11	305.9	15.1	302.7	14.4	3.2	3.95***
Nonacademic						
Grade 5	254.9	8.3	256.4	7.2	-1.5	2.21*
Grade 7	271.9	11.2	272.2	9.2	-0.3	0.34
Grade 9	285.7	14.8	286.7	13.6	-1.0	0.81
Grade 11	291.8	15.8	289.1	14.8	2.7	2.03*

*Significant at the 5 per cent level.

**Significant at the 1 per cent level.

***Significant at the 0.1 per cent level.


290
285
280
275
270
265
260
255
250
245
MEAN SCALE SCORES
5
7
9
11
GRADE
Academic Males
Academic Females
Nonacademic Males
Nonacademic Females


310
305
300
295
290
285
280
275
270
265
260
255
250
MEAN SCALE SCORES
5
7
9
11
GRADE
Academic Males
Academic Females
Nonacademic Males
Nonacademic Females

3. As early as the fifth grade, students who in high school enrolled in academic (college preparatory) programs had higher mean STEP Math and SCAT-Q scores than nonacademic students, even when classroom exposure to mathematics instruction is held roughly constant. From grade 5 to grade 9 the two curriculum groups diverge slightly, but from grade 9 to grade 11 the slopes of the trend lines are similar.

4. In STEP Math, males and females are equal at the grade 5 level but the males have successively higher mean scores at subsequent grade levels. For the academic group the sex differences are statistically significant at grades 7, 9, and 11. For the nonacademic group the difference is statistically reliable only at the grade 11 level.

5. Also in SCAT-Q the sex differences tend to increase with age. For the academic group the difference between males and females is statistically significant at the grade 11 level. In the nonacademic group, females obtained significantly higher scores than males at the grade 5 level, but in grade 11 the difference is in the opposite direction.

The absolute level and change in mathematics achievement observed here was fairly similar to that reported by Droege (1966) in a comparable study. In that large-sample study, involving the General Aptitude Test Battery, the girls had slightly higher mean scores in "numerical ptitude" at the grade 9 level, but by grade 12 the boys had pulled slightly ahead.

BEQ. We have noted increasing sex differences with age in quantitative ability and mathematics achievement. To find out if there existed corresponding patterns of sex differences in students' interest in mathematics, their views about the usefulness of learning mathematics, activities outside the classroom and so on, information was collected from the BEQ. Eight questions were examined, and the results are shown in Table 3. From the results

Table 3

Male-Female Differences in Percentage Giving Certain

Questionnaire Responses, by Curriculum

(N=632 males and 688 females in academic programs, and 249 males and 290 females in nonacademic programs)

Q.1: How many books on science have you read during the last two years?

A: 1 or more

	Academic				Nonacademic			
	Male	Female	Diff.	X^2	Male	Female	Diff.	X^2
Grade 7	28.6	24.4	4.2	2.99	23.7	19.6	4.1	1.33
Grade 9	35.7	19.1	16.6	46.00**	22.0	14.1	7.9	5.73*
Grade 11	29.6	11.0	18.6	71.48**	16.5	6.9	9.6	12.29**

Q.2: How often, on the average, have you read scientific magazines?

A: Occasionally or regularly

	Male	Female	Diff.	X^2	Male	Female	Diff.	X^2
Grade 7	35.5	34.8	0.7	0.07	18.1	19.4	-1.3	0.15
Grade 9	45.4	41.3	4.1	2.26	23.3	15.6	7.7	5.13*
Grade 11	49.0	37.8	11.2	16.84**	24.9	15.2	9.7	7.97**

Q.3: Were the math courses boring or interesting to you?

A_1: Boring

	Male	Female	Diff.	X^2	Male	Female	Diff.	X^2
Grade 7	10.1	11.2	-1.1	0.42	5.7	6.9	-1.2	0.32
Grade 9	11.8	14.2	-2.4	1.67	10.9	14.8	-3.9	1.80
Grade 11	6.0	17.9	-11.9	43.54**	20.9	26.2	-5.3	2.08

A_2: Interesting

	Male	Female	Diff.	X^2	Male	Female	Diff.	X^2
Grade 7	57.6	56.1	1.5	.30	47.4	50.7	-3.3	.58
Grade 9	70.1	62.2	7.9	9.16**	56.6	53.0	3.6	.70
Grade 11	64.5	53.0	11.5	17.94**	48.6	40.8	7.8	3.30*

Table 3 (Continued)

Q.4: Do you think the math courses will be useful in helping you earn a living?

A: Useful

	Academic				Nonacademic			
	Male	Female	Diff.	X^2	Male	Female	Diff.	X^2
Grade 7	76.0	74.3	1.7	0.51	69.8	67.6	2.2	0.30
Grade 9	86.7	77.3	9.4	19.55**	79.0	73.5	5.5	2.22
Grade 11	81.4	65.7	15.7	41.41**	72.7	59.0	13.7	11.10**

Q.5: How often have you talked about science with your friends?

A: Occasionally or frequently

Grade 7	50.1	48.8	1.3	0.22	44.2	36.5	7.7	3.31
Grade 9	62.4	43.0	19.4	49.69**	44.9	35.1	9.8	5.38*
Grade 11	53.9	31.9	22.0	65.27**	34.9	23.8	11.1	8.03**

Q.6: How often have you talked about science with your parents?

A: Occasionally or frequently

Grade 7	47.9	48.2	-0.3	.01	34.5	34.5	0.0	0.00
Grade 9	56.4	41.5	14.9	29.27**	36.9	27.9	9.0	4.98*
Grade 11	40.6	24.9	15.7	37.06**	28.5	14.8	13.7	15.09**

Q.7: How does your mother feel about your continuing your education beyond high school?[a]

A: Strongly favors it

Grade 9	90.2	86.2	4.0	5.03*	61.0	61.7	-0.7	0.03
Grade 11	91.3	83.9	7.5	16.43**	58.2	52.8	5.4	1.58

Q.8: How does your father feel about your continuing your education beyond high school?[a]

A: Strongly favors it

Grade 9	86.6	85.6	1.0	0.27	59.8	57.9	1.9	0.20
Grade 11	88.6	82.6	6.0	9.55**	59.4	52.1	7.3	2.89

[a]These questions were not included in the grade 7 BEQ given in 1963.

*Significant at the 5 per cent level.

**Significant at the 1 per cent level.

obtained the following conclusions can be drawn about the academic as well as the nonacademic group:

1. Reading books on science and scientific magazines is more frequent among males.

2. More males are interested in mathematics and more females are of the opinion that the math courses were boring to them.

3. More males think the math courses will be useful in helping them earn a living.

4. Talking about science with friends and parents is more frequent among males.

5. Parents more frequently favor a continuing of education beyond high school for their sons.

6. The most interesting finding is perhaps that all the obtained differences between males and females were generally negligible at the grade 7 level but increase with age.

7. As with the differences in achievement and ability the differences between males and females in the academic curriculum are generally greater than in the nonacademic curriculum.

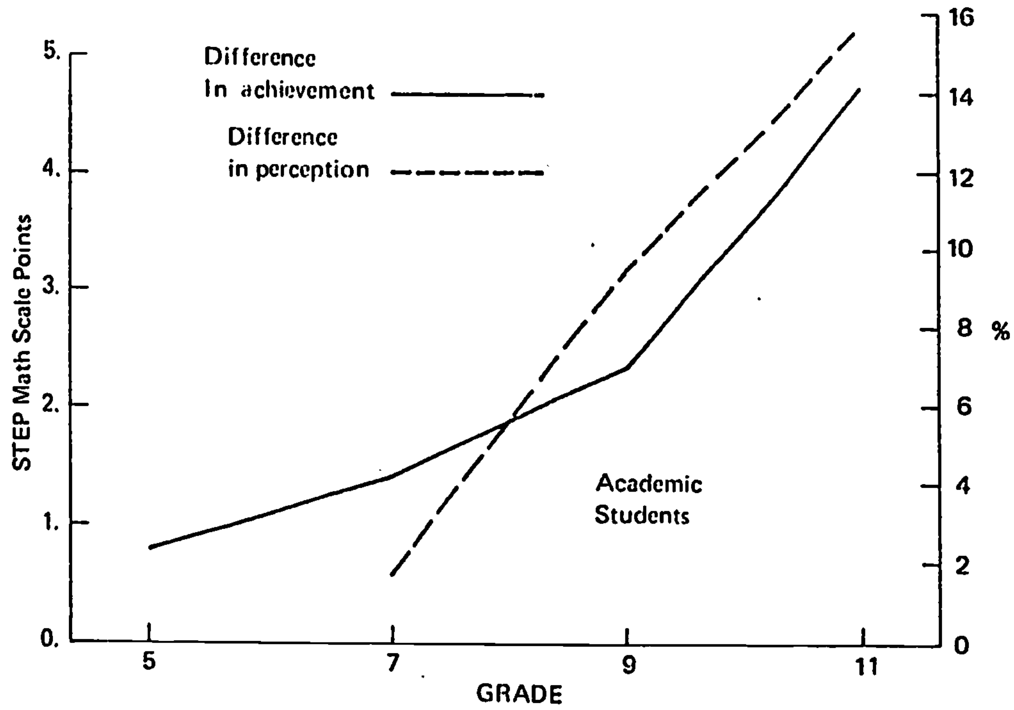
Discussion and Summary

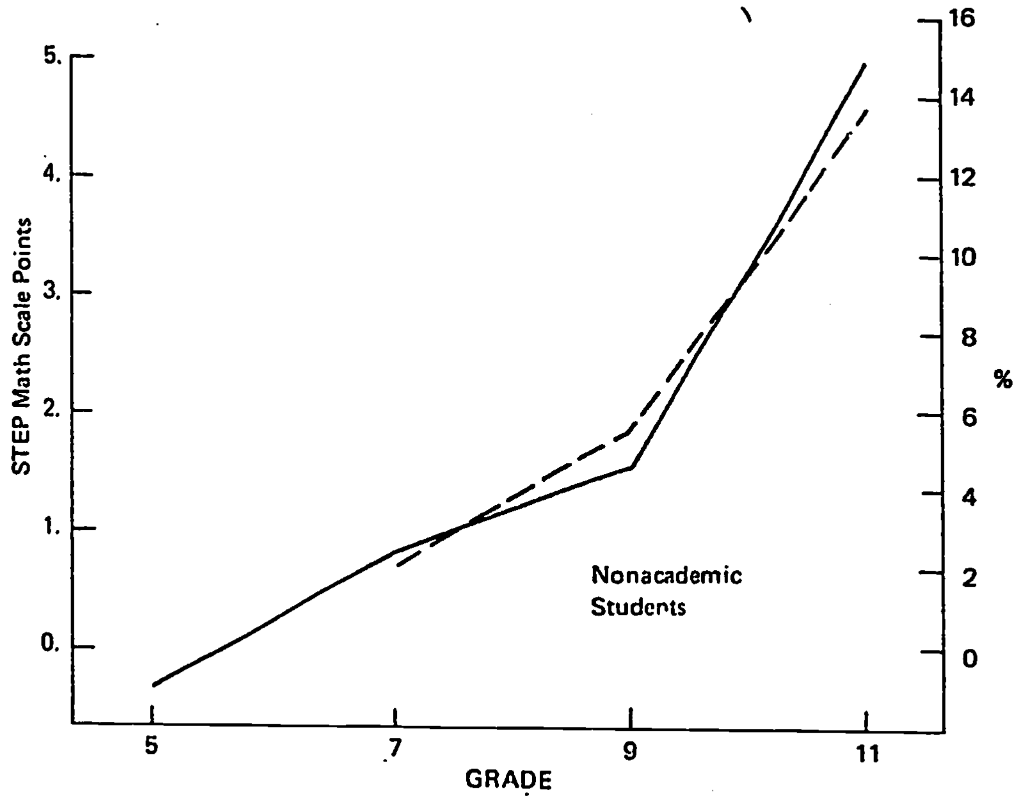
The main purpose of the present investigation was to study sex differences in mathematics achievement and how those differences change with age when curriculum identification and the amount of training in math are held roughly constant and when matched longitudinal samples are used. The results obtained reveal no sex differences in mathematics achievement at the grade 5 level.

This absence of difference is consistent with the hypothesis of sex-typed interests if it is assumed that most such interests emerge during adolescence. At subsequent grade levels (grades 7, 9, and 11) males have higher mean scores than females and the differences between the sexes increase with age (Figure 1). This conclusion holds for students enrolled in college preparatory programs (the academic group) as well as those who did not enroll in such programs (the nonacademic group), even though membership in all samples was restricted to students who were enrolled in mathematics during the periods in question.

Furthermore, the growing differences in math achievement between the males and females does, as predicted, take place in concert with increasing differences in interest. As the boys' interest in science increases relative to the girls' their achievement in mathematics increases relative to that of the girls. Figure 3 shows the relationship. In the variables plotted there is the difference between the boys and girls in math achievement and the difference between the boys and the girls in the percentage who perceived mathematics as useful in earning a living. (In each case the girls' statistics were subtracted from the boys'.) The congruence between the trend lines is striking. As far as Figure 3 is concerned, it would be more accurate to refer to the relationship between achievement and "anticipated utility." This is, no doubt, an ingredient of sex-typed interests but since we have no reason to focus only on anticipated utility we will continue to refer to "sex-type interests."

The real question in interpreting the congruence is the direction of causality. Aiken (1970, p. 558) points out that "the relationship between


Difference
In achievement
Difference
in perception
STEP Math Scale Points
5.
4.
3.
2.
1.
0.
16
14
12
10
8
6
4
2
0
%
Academic
Students
5
7
9
11
GRADE


STEP Math Scale Points
5.
4.
3.
2.
1.
0.
16
14
12
10
8
6
4
2
0
%
Nonacademic
Students
5
7
9
11
GRADE

attitudes and performance is certainly the consequence of a reciprocal influence, in that attitudes affect achievement and achievement in turn affects attitudes (see Neale, 1969). This dynamic interaction between attitudes and behavior has received a great deal of attention in the recent social-psychological literature (see Festinger, et al., 1964)." In other words, greater achievement results from an increase in interest and greater interest results from greater achievement. The exact nature of the relationship cannot, unfortunately, be investigated by means of the data at hand or, for that matter, from any data the authors can imagine. The problem is that the interaction between interest and achievement is probably instantaneous. One word of positive feedback from a respected teacher, guidance counselor, or peer and the student is immediately more interested in mathematics and, as a consequence, immediately more able in performing mathematically. When this process is examined at two-year intervals one can only hope to find evidence of a close correspondence between changes in interest and changes in ability, which is what was observed in this study.

Thus, although we cannot assert that sex differences in mathematics achievement <u>result from</u> sex-typed interests, we can say that the data from this study are not inconsistent with the hypothesis. If nothing else the data indicate there is a close relationship between a student's perception of mathematics and his performance in it. We see no need to hypothesize physiological or psychoanalytic explanations for the disparity in mathematical achievement between the sexes.

References

Aiken, L. R., Jr. Attitudes toward mathematics. Review of Educational Research, 1970, 40, 551-596,

Alpert, R., Stellwagen, G., & Becker, D. Psychological factors in mathematics education. School Mathematics Study Group Newsletter, 1963, No. 15, 17-24.

Anastasi, A. Differential psychology. (3rd ed.) New York: Macmillan, 1958.

Anderson, S. B., & Maier, M. H. 34,000 pupils and how they grow. Journal of Teacher Education, 1963, 14, 212-216.

Anttonen, R. G. A longitudinal study in mathematics attitude. Journal of Educational Research, 1969, 62, 467-471.

Carey, G. L. Sex differences in problem-solving performance as a function of attitude differences. Journal of Abnormal and Social Psychology, 1958, 56, 256-260.

Droege, R. C. Sex differences in aptitude maturation during high school. Paper presented at annual meeting of American Psychological Association, New York, 1966.

Festinger, L., et al. Conflict, decision, and dissonance. Stanford: Stanford University Press, 1964.

Ghiselli, E. E. Theory of psychological measurement. New York: McGraw-Hill, 1964.

Hilton, T. L. A study of intellectual growth and vocational development. Final Report, Project No. 6-1830 (Grant No. OEG-1-6-061830-0650). U. S. Office of Education, March 1971.

Hilton, T. L., & Myers, A. E. Personal background, experience and school achievement: An investigation of the contribution of questionnaire data to academic prediction. Journal of Educational Measurement, 1967, 4, 69-80.

Hilton, T. L., & Patrick, C. Cross-sectional versus longitudinal data: An empirical comparison of mean differences in academic growth. Journal of Educational Measurement, 1970, 1, 15-24.

Husén, T. (Ed.) International study of achievement in mathematics: A comparison of twelve countries. New York: Wiley, 1967.

Lindgren, H. C., Silva, I., Faraco, I., & Da Rocha, N. S. Attitudes towards problem solving as a function of success in arithmetic in Brazilian elementary schools. Journal of Educational Research, 1964, 58, 44-45.

Maccoby, E. E. (Ed.) The development of sex differences. Stanford: Stanford University Press, 1966.

Milton, G. A. The effects of sex-role identification upon problem-solving skill. Journal of Abnormal and Social Psychology, 1957, 55, 208-212.

Neale, D. C. The role of attitudes in learning mathematics. Arithmetic Teacher, 1969, 16, 631-640.

Tyler, L. E. The psychology of human differences. (3rd ed.) New York: Appleton-Century-Crofts, 1965.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. Psychology differentiation: Studies of development. New York: Wiley, 1962.